

Věc: **Zpráva z pracovní cesty**
Pracovník: **Pavčina Rouhová**
Akce: **Seminář krajské digitalizace**
Datum konání: **25.-26. 3. 2015**
Místo konání: **Moravská zemská knihovna, Brno**

25. 3. 2015

Stav digitalizace pro Krajské knihovny (KK) k 25. 3. 2015

- viz. tabulka - <http://goo.gl/Eay2fW>
- **Jihočeský kraj:**
 - aktuální verze K5 5.1, vloženo cca. 1,5 milionu stran, projekt digitalizace regionálních periodik - 360 000 stran, cca. 150 000 stran staré tisky z fondu JVK, cca. 1mil. stran periodik z K3 NKP - převedeno do JP2 formátu
 - LTP úložiště
 - digitalizační jednotka od ledna
 - 3 plné úvazky
 - centrum na KÚ
 - OCR vysoutěžili na fakturu
- **Plzeňský kraj:**
 - stále stejný stav v případě digitalizace z IOP, tj. zatím nedigitalizují
 - problém s výběrkem, odvolávají se dodavatelé
- **Karlovarský kraj:**
 - zdigitalizováno asi 200 tisíc stran a udržují
 - centrum na KÚ
- **Středočeský kraj:**
 - Něco mají digitalizováno, nahrají do Kraméria
 - zatím něco externě
- **Ústecký kraj:**
 - první fáze naplnění indikátorů projektu „Rozvoj služeb eGovernmentu v Ústeckém kraji“ byla úspěšně dokončena, nyní pokračuje druhá fáze ve stejném rozsahu, datum ukončení druhé fáze je stanoven na 31.7.2015
 - centrum na KÚ
 - dobrovolníci a důchodci
- **Liberecký kraj:**
 - Z KVKLI nikdo nedorazil
- **Královéhradecký kraj:**
 - digitalizace službou - formálně skončila k 20.01.2015, zdig. cca 500.000 stran, dodavatelem byla firma A-scan, subdodavatel pro metadata archlife/Incad, probíhají kontroly dat, něco reklamují, větší část už mají v K5, teď řeší opravy (špatné řazení ročníků a čísel – popis dle pravidel NDK vs zobrazování v K5)
 - krajská digitalizační jednotka – v prostorách SVK HK, od 06/2014 běží v ostrém provozu
 - skener Bookeye3 A2 + SW BCS-2, metadatový popis v NDK verzi ProArcu
 - od 1.3. běží udržitelnost
 - úvazky placené z projektu k 28.02. skončily, ale od 01.03. pro práci na KDJ dostali 1 úvazek do stavu SVK + na 03-08/2015 mají ještě cca 2,5 úvazku z projektu ÚP Odborné praxe pro mladé do 30 let

- **Pardubický kraj:**
 - externě probíhá výběrové řízení
 - mají cca 50 tis. (nebo 500)
- **Kraj Vysočina:**
 - 2 plné úvazky
 - digitalizují periodika a kompletují
 - digitalizují na službu
 - výběrové řízení na Portál digitalizace
 - 3D digitalizace
- **Olomoucký kraj:**
 - data z externí digitalizace zpracována, nahrána do Krameria
 - probíhá průběžná vlastní digitalizace (k dispozici tři skenery) a zpracování
 - využívají Sirius, MEditor (také ProArc)
 - data uchovávají v repozitáři (dlouhodobá ochrana)
 - do Krameria postupně začleňují i dříve zpracované dokumenty a replikují z K3 do K5
 - nyní obsah téměř 1 500 000 stran
 - 2 úvazky
- **Jihomoravský kraj:**
 - digitalizují externě pro několik muzeí – vystaví MZK (k5)
- **Moravskoslezský kraj:**
 - 1 ½ úvazku
 - dokontrolovávají
 - digitalizují vlastní fond i fond ostatních
- **Zlínský kraj:**
 - nic nového

Metadatové standardy / Jarek Kvasnica, Iveta Lodrová (NKP)

Změny standardů z října 2014:

- periodika 1.5.1 – oprava chyb
- monografie 1.1.2 – text, mapy, vícesvazky
- povinné urn:nbn
- název PSP balíčku dle UUID nebo urn:nbn

Identifikátory:

- urn:nbn povinně
- neplatné identifikátory

Autorsko-právní metadata:

- copyright MD
- dědičnost
- jednoduchý komplexní zápis
- nepovinný ...

Info XML:

- větší důraz na soubor
- povinnost elementů
- možnost zapsat ...

Nová pravidla RDA:

- drobné změny (ale zásadní) budou i ve standardech:
 - návěští
 - nakladatelské údaje
 - ODD (336, 337, 338)
- chystají se změny v transformačních šablonách, aby knihovny mohly začít generovat nové záznamy dle RDA, budou dvojí záznamy (staré dle AACR2 a nové dle RDA), MODS zůstává stále stejný, jde o změnu polí v šablonách pro převod z MARC21 do MODS, v hlavičce záznamu MARC (Leader, pole 18) bude informace o tom, jestli je záznam katalogizován dle RDA nebo AACR2, podle toho se aplikuje vhodná šablona na správný převod polí, v MODS zůstanou pole stejná, jen se metadata u RDA budou brát z jiných polí
- aktualizace pravidel popisu pro editaci (spolupráce NK a MZK) i metadatové specifikace
- termín: cca konec dubna 2015
- *diskuze digitalizátorů o tom, co je to vlastně RDA – fakt sranda ☺*

PSP balíček

- info.xml soubor – info o vzniku a obsahu balíčku, důležité je pojmenování balíčku, už v hlavičce jde poznat, jestli jde o periodikum nebo monografii podle čísla verze metadatové specifikace (1.5.1 – periodika, 1.1.2 – monografie)
- soubor MD5 info.xml jako jediné nemá, protože by docházelo k součtu stále dokola
- validátor metadat – byla snaha začít vyvíjet nový a lepší, návrh do VISK3, ale nevyšlo to
- Elsysť má taky validátor – lepší, ale dost drahý
- md5 soubor – všechny součty kromě info.xml souboru
- mets.xml – hlavní METS soubor
- mods – na úrovni čísla je MA pro místo vydání a nakladatele – není potřeba
- md5 u každého souboru i tady – je to pro jistotu kvůli kontrole; které číslo je v procesu kontroly použito, je nahodilé, čím víc md5 souborů, tím líp :)
- amd_mets.xml
- alto.xml – u každého slova udává jeho pozici, kde se na straně nachází, umožňuje to, že ve fulltextovém vyhledávání vyhledávač „skočí“ na stránce přímo na hledané slovo (zobrazí ho v rámečku)
- NDK – komplexní validátor – vývoj
- mets :
 - mods – jsou bibliografická metadata
 - strukturální (fyzická) mapa - reprezentace té které strany
 - logická mapa - pracuje s částí popisu
 - order label – název stránky
 - vedlejší mets – administrativní a technická metadata
- NDK standardy:
 - mix – pro podrobný popis obrázků
 - premis – pro dlouhodobou ochranu
- květen 2014:
 - platí, že balíčky se pojmenovávají dle URN:NBN

Kramerius – webová a mobilní aplikace, editor, administrace

Webová aplikace:

- klient Incadu – nic moc, je nedokonalá, připomínkováná a uživatelsky nepřívětivá
- oproti tomu aplikace od Honzy Rychtáře, uvidí se, jak s tím MZK (a ČDK) naloží
- mobilní verze K5 – je třeba mít min. K5.02 + logo (a vybrané dokumenty)
- digitální knihovna na YouTube
- Kramerius je na Obchod Play (Android)
- *tiff se používá teprve poslední 3 roky, od NDK, dříve se nepoužíval, protože byl příliš velký, používal se jpeg*
- prolinky na plné texty v Alephu na Krameria (pole 001 – NDK)
- **oficiální Kramerius 5:** ke shlédnutí např. na:
 - <http://cdk-test.lib.cas.cz/client/index.vm>
 - <http://kramerius.mzk.cz/client/index.vm>
- připomínky na Github:
 - <https://github.com/ceskaexpedice/kramerius/issues>
- zde jsou vyfiltrované ty, které se týkají webového klienta K5:
 - <https://github.com/ceskaexpedice/kramerius/labels/K5Client>
- **neoficiální Kramerius 5:**
 - <http://janrychtar.cz/kramerius>
 - aby se načítaly data, bude potřeba nainstalovat Kramerius 5.1 + ještě patch, který řeší přístup prohlížeče na server

Aplikace pro Android:

- tablety v MZK – skrytá wifina, dá se dostat i k neveřejným dokumentům v budově
- když uživatel opustí budovu (s tabletem), skrytá wifina už nefunguje, takže k neveřejným dokumentům se nedostane
- když si uživatel něco stáhne, v kódu toho dokumentu je informace, za jakých podmínek lze obsah prohlížet (nebo v rámci jaké wifi), takže mimo MZK si stažené obrázky neprohlédne
- <https://play.google.com/store/apps/details?id=cz.mzk.kramerius.app>
- **podmínky pro zařazení digitální knihovny do aplikace:**
 - Kramerius 5.0.2 (nejnižší možná verze API), logo, vybrat dokumenty na domovskou stránku knihovny a napsat nám email
- **připomínky k aplikaci:**
 - psát na developer@mzk.cz nebo vytvořit issue na Githubu
 - <https://github.com/moravianlibrary/kramerius-for-android/issues>

Aplikace pro iOS

- vytváří programátor z FI MU v rámci diplomové práce, částečně placený z VISKU 3
- v MZK dostali dotaci z VISKU 3 na vývoj mobilních aplikací Krameria
- plánuje se přehrávání zvuku v aplikaci i na pozadí, když si chce člověk číst booklet nebo prohlížet jiné obrázky patřící ke kolekci
- v rámci dotace se přidá i swipování

Dotazy k administraci, editoru, Krameriovi

- kapitoly v knize – model internal part, není možné vytvářet podkapitoly kapitol, není na to v Krameriu model, je to asi jako snaha vložit přílohu (která má model čísla) do čísla
- někomu nejdou nahrát do MEditoru obrázky v tiff, ale v jpg jdou ... obrátit se na Martina Rumánka (martin.rumanek@mzk.cz)
- dříve se uchovávaly obrázky v jpg, velká část archivu NK má obrázky v jpg, potom se začal uchovávat tiff, djvu, jpeg 2000
- vyhledávání pomocí víceslovných klíčových slov – někomu nefunguje, v MZK to funguje, možná, že to vyladili v nové verzi a v těch starších to nejde
- vytváření virtuálních sbírek – nastavení IP adres, na kterých jdou konkrétní dokumenty zpřístupnit, nelze nastavit pro celou sbírku, návrh na vývoj?
- zveřejňování periodik – MZK 70 let od vydání, NK má 115 let, záleží na právním výkladu, nemohly by se instituce domluvit ... asi ne
- automatizované odkazy z Alephu do Krameria – je na to skript, dá se to udělat přes oai protokol, kvůli skriptu na import do Alephu lze kontaktovat Tomáše Prachaře tomas.prachar@mzk.cz, skript rád poskytne
- vytvářel někdo pdf soubory do tifu?
- PDF/A v Krameriu – zobrazí se (v Chrome a Firefox je plugin přímo v prohlížeči, jinak se navrhne stažení dokumentu)
- Github – K4 tools: <https://github.com/moravianlibrary/K4-tools>
- MZK - bude se vytvářet skript na dávkové přidání odkazů do Alephu z Krameria u digitálních dokumentů NDK NKP, které v Alephu ještě nemají (možná pomocí ČČNB, možná přes souborný katalog, ještě to nezačali řešit prakticky)

26. 3. 2015

Diskuze / Petr Žabička

- CPK
- OCR fraktury – NK – bude rozpoznávat obálky knih, licence na OCR, spolupráce se Scanservisem, byla snaha domluvat se s p. Tichým, to ale nedopadlo

ProArc systém (prezentace z KNAV) / Martina Nezbedová

- KNAV používá:
 - Krajská knihovna ve Zlíně
 - Studijní a vědecká knihovna v HK
 - MKP
 - ...
- spokojenost se systémem
- představení systému – PSP balíčky, dle NDK standardů, založeno na systému Kramerius (Fedora, PostgreSQL, Kakadu atd.)
- typy dokumentů – modely NDK (monografie, periodika atd.)
- MZK – po vyřešení Lidovek to zkusí sami nainstalovat a vyzkoušet
- <https://code.google.com/p/archivacni-system>
- registr digitalizace – přihlášení přes čarový kód

LTP / Zdeněk Hruška

Projekt LTP Pilot

- financováno z Fondu rozvoje sítě CESNET
- září 2014 - říjen 2015 (14 měsíců)
- UVT FI MU, MZK, externí – A. Fojtů, Hutař, Melichar
- open source LTP systém Archivematika (AV)
- www.archivematica.org
- testování Archivematicy (nyní verze 1.3) – kanadská firma Artefactual Systems (financováno i z jiných knihoven, které si zaplatí přidání nějaké funkce)
- novinka - Archivematica v cloudu
- běží na Ubuntu 12.04 LTS
- open source nástroje, složení z mikroslužeb (microservices), workflow, ingest funguje relativně dobře (ingest různých typů dat), širší kontrola dokumentů (kopie, kde jsou uloženy atp.) ale není vyřešena – potřeba propojit s jiným systémem (např. DuraCloud), uuid, md5, identifikace formátů (např. Fido)
- nemá svoje úložiště - data se ukládají do Ubuntu, počítá s tím, že svoje úložiště bude mít instituce vyřešené mimo AM
- Finsko – používají Archivematicu jako správce workflow, defaultní mikroslužby vyhodili a udělali si nové, k tomu se přiklání i MZK – upravit systém dle potřeb
- testování – výkon HW + SW (zátěžové testy, scénáře), snaha napojit na úložiště CESNETU (pro knihovny využití zdarma) – možnost, že by každá instituce měla svou Archivematicu a všechna data byla uložena tam
- kompatibilita s normami OAIS a ISO 16363 (jsou tam nějaké mouchy)
- strategie dlouhodobé ochrany – normalizace + uchovává i původní formát
- výstup – manuál, co systém umí a neumí, dokumentace, jak ho nainstalovat a pracovat s ním, články ohledně Archivematicy, české překlady
- testování na – <https://archivematica.ics.muni.cz>
- archivematica indexuje všechny textové informace
- LTP portál – LTP obecně, snaha o rozvoj LTP v ČR, podpora komunity
- MZK - pokusí se KK informovat o dalších aktivitách týkajících se LTP
- <https://docs.google.com/presentation/d/1oibzNEZ9ea9TfaNzE9mYa63zbFgg0ZK3Lh43qbhKd28/edit?usp=sharing>

LTP projekt na KISKu

- APLS – dvousemestrální projekt studentů KISKu
- potřeba osvěty a vzdělání v oblasti LTP
- snaha o vybudování komunity kolem LTP - zaměřeno na knihovny ČR
- výstupy:
 - newsletter 1x měsíčně (články o LTP, novinky, seriál článků „úvod do LTP“)
 - webináře k LTP
- <http://ltp.knihovna.cz>
- Facebook: <https://www.facebook.com/groups/1401935900117180/?fref=ts>
- Twitter: https://twitter.com/ltp_cz
- https://docs.google.com/presentation/d/1kdQLXv08oGCTUkXx6bot0R7a0VgNDMG5kvA_njFcJlg/edit?usp=sharing



DSA - Data Seal of Approval

- základní certifikace dlouhodobého úložiště
- směrnice pro repozitáře (český překlad směrnice DSA)
- certifikované DSA repozitáře (UK):
 - <http://dsa.cuni.cz/>
- self-audit

Gramodesky, VNF / Filip Šír

- oddělení digitalizace zvukových dokumentů MZK - malé oddělení, ale řeší velké věci :)
- první takové oddělení v ČR v knihovnách
- zvukové nahrávky do Krameria - jenom skrz Meditor
- projekt NAKI II.
- možná spolupráce se Slovenskem
- <https://www.narodnifonoteka.cz/>
- <projekt.narodnifonoteka.cz>